

REFLEXIÓN Y SOCIALIZACIÓN

ARTÍCULOS DE INVESTIGACIÓN,
ORIGINALES E INÉDITOS, Y OTROS TIPOS
DE COLABORACIONES CIENTÍFICAS

EL IMPACTO DE LA INTELIGENCIA ARTIFICIAL EN LA JUSTICIA

THE IMPACT OF ARTIFICIAL INTELLIGENCE ON JUSTICE

■ **DRA.C. EMILIA MARÍA SANTANA RAMOS**

Profesora de Filosofía del Derecho,
Universidad de Las Palmas de Gran Canaria, España

<https://orcid.org/0000-0002-7390-4065>

emilia.santana@ulpgc.es

Resumen

La actualidad evidencia cómo la inteligencia artificial ha irrumpido en la órbita planetaria e impactado en todos los sectores sociales; evidentemente, el Derecho no es la excepción. Las posibilidades que ella ofrece, en relación con la capacidad para procesar y almacenar datos, abre un abanico de probabilidades en la administración de justicia. Sin embargo, su incorporación en la toma de decisiones judiciales está sujeta a discusiones de gran calado ético y jurídico. El presente trabajo analiza los beneficios y riesgos de tal uso. Si bien es cierto que la inteligencia artificial promete mejorar la eficiencia y optimizar los procesos, es preocupante, realmente, reconocerla como una herramienta capaz de dictar decisiones judiciales. Las preguntas obligadas se centrarían en la responsabilidad en caso de un fallo erróneo, la garantía de la proporcionalidad, la salvaguarda de los derechos fundamentales y la limitación de los sesgos en la toma de decisiones automatizadas, con respecto a la equidad. Evidentemente, estas cuestiones no son baladí y, por tanto, no han de ser ignoradas.

Palabras clave: Inteligencia artificial; administración de justicia; decisiones judiciales; legitimidad de la justicia; derechos fundamentales.

Abstract

The current situation shows how artificial intelligence has burst into the planetary orbit and impacted on all social sectors; obviously, law is no exception. The possibilities it offers, in relation to the capacity to process and store data, opens up a range of probabilities in the administration of justice. However, its incorporation in judicial decision making is subject to profound ethical and legal discussions. This paper analyses the benefits and risks of such use. While it is true that AI promises to improve efficiency and optimise processes, it is worrying to recognise it as a tool capable of making judicial decisions. The necessary questions would focus on liability in the event of an erroneous ruling, ensuring proportionality, safeguarding fundamental rights and limiting bias in automated decision making, with respect to fairness. Clearly, these issues are not trivial and therefore should not be ignored.

Keywords: Artificial intelligence; administration of justice; judicial decisions; legitimacy of justice; fundamental rights.

Sumario

I. Impacto de la inteligencia artificial en el Derecho; II. La inteligencia artificial en la toma de decisiones judiciales; III. El mito de la neutralidad: cuando el sesgo se esconde tras la objetividad; IV. Conclusiones; V. Referencias.

I. IMPACTO DE LA INTELIGENCIA ARTIFICIAL EN EL DERECHO

La inteligencia artificial (IA) se configura en un contexto teórico y conceptual en el que se reconoce que esta tiene como objetivo el desarrollo de sistemas computacionales que cuentan con capacidad para actuar en espacios en los que se requiere de la inteligencia humana. Se hace referencia al razonamiento, el aprendizaje, la percepción y la toma de decisiones, etc. Así la reconoce Cotino (2017), cuando defiende que ella «permite generar patrones dinámicos de

tendencias de futuro: la predictibilidad y el apoyo en la toma de decisiones» (p. 133). De igual manera, otros autores la entienden como «la capacidad de las máquinas para imitar y ejecutar tareas que requieren inteligencia humana» (Farfán, 2023, p. 175).

Esta problemática recuerda el debate planteado por Hart (1961), quien sostenía que el Derecho es un sistema de reglas primarias y secundarias. En el pensamiento del autor, podría entenderse la aplicación de las reglas primarias (obligaciones legales) como un intento de automatizar, pero sin la capacidad de interpretar las normas, atendiendo a otras cuestiones en el caso concreto (pp. 102-103).

En el contexto actual, no es discutible el impacto provocado por la IA y la importancia en su empleo y manejo, pues la capacidad para replicar procesos cognitivos humanos, como el razonamiento y la toma de decisiones, la convierte en una herramienta con el poder de proyectarse en los diferentes marcos, como los económicos, políticos, sociales y jurídicos.

Desde algunos sectores proponen identificarla, atendiendo a la capacidad algorítmica que le permite un aprendizaje por el cual es capaz de instruirse a partir de datos; ese conocimiento pudiera destinarse a la adopción de decisiones, cual si se tratase de un ser humano. A diferencia de las personas, los sistemas de IA pueden operar sin descanso y analizar grandes volúmenes de información simultáneamente, lo que les permite ejecutar tareas con mayor eficiencia y con una proporción de errores considerablemente menor. Para Rouhiainen (2018), son destacables las características que la identifican. En este sentido, apunta a la capacidad que tiene de almacenamiento de datos y de aprendizaje para adoptar decisiones, lo que significa que una máquina puede convertirse en un actor con posibilidad de adoptar decisiones concretas en un ámbito específico (p. 17) o, lo que es lo mismo, «trasladar el modo de pensamiento y razonamiento humanos a la computación. Particularmente se propone desarrollar en la máquina la capacidad de percepción, el aprendizaje a partir de la experiencia, relacionar conocimiento incluyendo conocimiento abstracto o conceptos y resolver problemas» (Alastruey, 2021, p. 183). Permitir la sustitución de una responsabilidad que, en principio, debería asumirse por una persona, por la de los sistemas inteligentes, debe llevar consigo un análisis cauteloso, sobre todo, en los marcos concretos en que esa situación se produzca.

Debido a las particularidades inherentes a la IA, algunos autores sostienen que, para comprenderla mejor, es necesario analizar tanto sus componentes como su alcance. De este modo, se hace menester precisar, desde un punto de vista conceptual, su significación. En la nomenclatura de la IA, el término *inteligencia* se refiere a la capacidad de razonar, aprender y tomar decisiones. Cuando se alude al concepto de artificial, aflora la falta de univocidad del término, lo que impide determinar, de forma nítida, la naturaleza de la IA y su relación con la inteligencia humana (Pazos y García, 2024, p. 8).

Esto permite preguntarse hasta qué punto uno estaría dispuesto a renunciar al raciocinio humano por la capacidad generativa de lo artificial. Este parecer pone de relieve que determinados elementos, pese a ser artificiales, pueden dar una apariencia de naturalidad. Por ejemplo, el césped artificial y el natural. No se duda de que ambos cumplen la misma función decorativa, pero no tienen las mismas propiedades. Lo que se quiere significar es que no basta con que una máquina simule el comportamiento humano, si no comprende ni experimenta la realidad de igual modo. Así como el césped artificial no respira ni interactúa con el entorno de la manera en que lo hace el natural, una IA, por avanzada que sea, no cuenta con sentimientos, no interpreta el mundo con subjetividad ni construye conocimientos desde la experiencia vivida.

Es por ello que, en este trabajo, se pretende abrir una reflexión sobre la importancia de la implementación de la IA en un espacio en el que entran en juego los derechos fundamentales (DF). Esto obliga a considerar que la IA posee una estructura basada en algoritmos, *software* y *hardware*, mientras que la respuesta humana proviene de redes neuronales y procesos racionales que integran conciencia, emoción y percepción sensorial, entre otros elementos.

Tales cuestiones cobran una importancia significativa cuando se trata de la salvaguarda de los DF en la interpretación y aplicación del Derecho. Imputar un significado a un enunciado normativo requiere tanto de una lógica humana como de una interpretación de valores, especialmente en lo relativo a DF como la libertad, la tutela judicial o la privacidad, entre otros.

Ahora bien, la realidad se impone: no se está ante una mera ilusión tecnológica, más bien, ante un sistema que aspira a replicar el raciocinio humano sin alcanzarlo realmente. El discurso sobre el remplazo

de la mente humana por la máquina no se puede decir que sea actual; ha estado presente en diferentes momentos históricos. Es muy acertado el pensamiento de Habermas (1948) cuando reconoce que «las realizaciones de la técnica, que como tales son irrenunciables, no podrán ser sustituidas por una naturaleza que despertara como sujeto» (p. 63). Aun asumiendo la postura de este autor y reconociendo que la técnica y la ciencia son irremplazables, por cuanto se configuran como un potencial en el desarrollo humano, en ningún caso aquellas pueden concebirse como un remplazamiento de la máquina sobre el humano ni en el servilismo del hombre hacia el artefacto.

Aunque exista la tentación de imaginar una alternativa en la cual la tecnología no domine la vida humana, ella solo sería posible si se modificara, por completo, la manera en que la especie interactúa con su entorno. Y, en ese contexto, no se puede olvidar que el constructo social se fundamenta en una racionalidad instrumental o, lo que es lo mismo, una técnica como medio para alcanzar fines específicos. Es indiscutible que un sistema computacional no debiera asumir el rol de una persona —y tampoco debiera hacerlo—, pues, con independencia de que se permita a los sistemas de la IA simular un comportamiento, ellos nunca podrán asumir la respuesta de la misma manera en que lo haría un ser pensante, ya que, evidentemente, la comprensión, en el análisis de la realidad, se percibe de diferente manera.

Este cambio de paradigma muestra cómo los constructos sociales, bajo el racionalismo técnico-científico, se configuran como una brújula de acción. Ello se demuestra cuando las sociedades se preocupan más por alcanzar una serie de beneficios y rentabilidad para optimizarlos. Desde este particular, se entiende que el avance tecnológico se centra más en el principio de utilidad, aunque tenga que superar los límites éticos o jurídicos. En efecto, se trata de implementar la IA con el claro objetivo de *optimizar*, sin tener en cuenta los derechos que entran en juego, como la privacidad, la libertad o seguridad personal. Esto lleva a plantearse el dilema entre la racionalidad y las decisiones generadas por la IA.

El impacto de esta racionalidad no se limita al ámbito productivo o científico. Sus principios se han trasladado al mundo social e influyen en la manera en que se estructuran las instituciones y las relaciones

humanas. Como consecuencia, se impone una lógica en la que todo se somete al cálculo, la planificación y la eficiencia, a la vez que se desplazan otras formas de pensamiento y decisión. Incluso, aspectos esenciales de la vida, como la naturaleza o la comunicación, son reinterpretados bajo esta lógica cuantificadora y pierden su dimensión cualitativa. Algunos autores ya advierten que, al carecer de sentido común, la IA no cuenta con la posibilidad de sostener un argumento sólido, con independencia de que sí puede replicar o simular silogismos formalmente válidos. «Lo que una máquina puede aportar a nivel argumentativo, es reproducir silogismos formalmente válidos, pero no sólidos. La veracidad que un juez debe determinar depende de características humanas que un sistema de IA no puede replicar» (Cabrera, 2024, p. 194).

No se está en presencia, únicamente, de avances tecnológicos, más bien, de una transformación de la sociedad y la manera en que se concibe la existencia, como resultado de la IA. Es decir, la evolución o involución a la que el ser humano está condenado, irremediablemente, radica en la eficiencia y la optimización. Así, la vida y sus interacciones corren el riesgo de ser reducidas a simples datos o procesos ajustables, y generar un mundo en el que la razón instrumental sustituya la reflexión crítica y la deliberación. Es interesante el aporte que ofrece García-Marzá (2024) cuando asevera que

el interés técnico busca el dominio, la superación de los límites impuestos por la naturaleza. Pero el éxito de sus resultados no tarda en aplicarse al mundo social, a la sociedad y a sus instituciones. En consecuencia, se prioriza una forma de pensar en la que todo se somete al cálculo y planificación [sic], donde incluso la vida, la naturaleza y la comunicación humana, quedan interpretadas bajo este cálculo de resultados. (p. 106)

No es extraño, por tanto, que la IA y la automatización sean consideradas no solo como herramientas tecnológicas, ya que representan la última expresión de un proceso más amplio de desplazamiento de la capacidad humana por sistemas técnico-administrativos. Son riesgos presentes en el pensamiento de Truyol (1991, p. 118), quien sugiere no olvidarlos, aun cuando reconoce los beneficios que llevan aparejados las tecnologías. Lo que, en realidad, resulta importante

es, precisamente, quién controla la IA y con qué propósito se implanta esta. Desde luego, se coincide con este parecer, ya que no solo se habla de innovación tecnológica, también se plantea la necesidad de garantizar que el avance de la IA continúe al servicio del ser humano y favorezca su desarrollo, evitando nuevas formas de dependencia. Desde esta perspectiva, más que discutir sus beneficios o su potencial, el debate se centra en la capacidad para orientarla, de manera que promueva el bienestar y la autonomía de las personas. Ahora bien, existe el riesgo —que no pocos temen— de que la tecnología deje de ser un medio y se convierta en un fin en sí misma, y el de que la voluntad humana se someta a dinámicas automatizadas o estructuras de control difíciles de revertir.

Se pone en cuestión la capacidad de la IA para simular procesos cognitivos, pero, también, el hecho de que se trata de un diseño concebido por seres humanos con el propósito de replicar ciertas funciones de la mente e, incluso, remplazarlas. Lejos de ser una abstracción teórica, la IA tiene aplicaciones tecnológicas concretas y consecuencias tangibles en múltiples ámbitos. Desde esta perspectiva, resulta imprescindible evaluar el papel y el impacto que ella genera en la transformación de los constructos sociales y jurídicos. Porque, indiscutiblemente, la proyección inmediata de la implementación de la IA redefine la manera en que se interactúa con la tecnología, al tiempo que obliga a replantear nociones como autonomía, dignidad, intimidad, libertad.... Desde luego, sería aberrante que la racionalidad instrumental pueda sustituir esferas que, por la propia naturaleza de los derechos en juego, no queden gobernadas por una lógica racional, porque, de lo contrario, en vez de reflexionar sobre la justicia, se haría un replanteamiento sobre qué es más óptimo en términos de eficiencia e, incluso, aquellos marcos en los que la eficiencia no debería ser el único criterio de decisión. Así lo defiende Otero (2023) cuando reconoce que «hay operaciones de inteligencia que precisan la utilización de características que sólo las personas tienen y nunca podrán tener las máquinas, por ejemplo, las emociones, la conciencia, la intuición etc.» (p. 47).

En el ámbito jurídico, la implementación de la IA sería un ejemplo palmario de lo antedicho, pues va más allá de la eficiencia o la automatización de procesos y cómo se debe gestionar la aplicabilidad de la norma en relación con la justicia y el Derecho. En este punto,

no se quiere manifestar una postura contraria a los beneficios referidos a la capacidad de la IA para la recopilación de datos y la agilización de respuestas. Sin embargo, el Derecho no puede reducirse a un ejercicio mecánico de aplicación de normas. El sentido de justicia no debe reducirse a un compendio de datos o patrones de actuación, ya que implica, también, la exigencia de una interpretación basada en principios que no siempre pueden ser reconocidos por un algoritmo. Los operadores jurídicos deben ser cautelosos en el manejo de la IA, a la hora de adoptar una decisión. Cobra sentido la preocupación que se despierta en la doctrina sobre la posibilidad de que los algoritmos reproduzcan sesgos estructurales que tengan como respuesta desigualdades o prácticas discriminatorias o, incluso, las refuercen. Son muchos los debates que se han abierto sobre la posibilidad de considerar legítima una decisión, basada en un sistema que carece de la capacidad de fundamentar la toma de decisiones que afectan DF de razonamiento jurídico. Desde luego, atendiendo a la dimensión axiológica del Derecho, la pregunta obligada sería si los algoritmos, por muy desarrollados que se encuentren, cuentan con la equidad, proporcionalidad o ponderación necesarias en la aplicación e interpretación del Derecho. El núcleo fundamental y problemático lo constituye el empleo de la IA en el proceso judicial. Desde una visión crítica, este tema obliga a repensar la labor del juez como garante de la justicia y la ponderación que caracterizan al Derecho, si una y otra son sustituidas por la IA.

II. LA INTELIGENCIA ARTIFICIAL EN LA TOMA DE DECISIONES JUDICIALES

Que la IA ha irrumpido en los constructos sociales es un hecho evidente. Con mayor frecuencia se observa cómo, en todos los sectores (económicos, sociales, sanitarios, etc.), ejerce una influencia más que significativa. En el contexto que se aborda, el que la IA encuentre acomodo en el ámbito jurídico, con la promesa de agilizar procesos, reducir la carga de trabajo de los tribunales y, en algunos casos, predecir decisiones judiciales, posiciona a los operadores jurídicos en una situación, cuanto menos, expectante.

Sin entrar en la disquisición sobre las fortalezas de la IA en términos de eficiencia, su implementación en las decisiones judiciales abre la puerta a reflexionar sobre los atributos del Derecho y la legitimidad

de la función judicial. Se trata, por tanto, de evaluar su eficacia operativa, y de cuestionarse hasta qué punto una máquina puede asumir el papel de garante de la justicia, sin poner en riesgo los principios fundamentales que sustentan el sistema jurídico, como la equidad, la imparcialidad y el respeto a los DF. La justicia no es más que el reflejo de un constructo social; por ello, no puede ser concebida como un valor estático. La realidad sobre la posibilidad de que la IA se convierta en una herramienta que, lejos de contribuir como un apoyo, se configure como un órgano activo, en la toma de decisiones judiciales, despierta un debate que no queda exento de contradicciones, sobre todo, en el papel del juez en el sistema de la administración de justicia. Pues, ¿hasta qué punto se puede confiar en que los sistemas computacionales desempeñen la labor tradicional del juez?

Hasta el momento, se ha defendido la IA como una socorrida herramienta de apoyo en la gestión de expedientes. Sin embargo, de manera inexorable, ella está revolucionando el concepto tradicional en la práctica jurídica, tanto en la interpretación del Derecho como en su aplicación, lo cual redimensiona la lógica algorítmica. Lo anterior obliga a meditar sobre el nuevo modelo de escritura jurídica que se enfrenta, en el que la argumentación jurídica tradicional es sustituida por la predicción automatizada o, al menos, condicionada por esta; por ende, ha de evaluarse la redefinición del valor o sentido de la justicia.

Algunos autores reconocen las fortalezas de la IA como instrumento de auxilio para optimizar la administración de justicia. En tal sentido, Cabrera (2024, p. 186) identifica dos algoritmos de evaluación de riesgos que analizan la probabilidad de reincidencia en Derecho penal, y los sistemas jurídicos expertos, centrados en la predicción de razonamientos interpretativos mediante el análisis de legislación y jurisprudencia.

Sin embargo, el verdadero punto de inflexión llegará cuando la IA no solo se configure como una herramienta auxiliar, más bien, cuando asuma un rol más determinante. La pregunta obligada no es si la IA podrá sustituir la labor del juez y cómo asegurar la protección de los derechos, la justicia y la independencia, además de prevenir la arbitrariedad. Desde una perspectiva garantista, Ferrajoli (2006) sostiene que el Derecho debe estructurarse en torno a la protección de los DF, mediante un sistema de garantías; considera como tal «cualquier técnica normativa de tutela de un derecho subjetivo» (p. 31). La implementación de la IA

en la justicia obliga a pensar en cómo pueden quedar garantizadas las decisiones automatizadas para que cuenten con los principios que inspira el debido proceso. El autor antes mencionado —coincidiendo con Dworkin en el reconocimiento de las reglas primarias y secundarias— muestra su preocupación por el sistema de garantías institucionales. En su línea de pensamiento, se entiende que el poder judicial debe actuar como un contrapeso que limite la arbitrariedad y asegure, por ende, que los jueces resuelvan conforme a una argumentación racional inspirada en el control de constitucionalidad.

En este marco, surge la preocupación de que los algoritmos, al no estar sujetos a una reflexión moral ni a un control judicial tradicional, puedan emitir fallos que vulneren las garantías fundamentales. Esto pone en cuestión si la IA es compatible con un modelo garantista de justicia.

En el razonamiento dworkiniano, el Derecho es concebido como un sistema de reglas que dependen de principios como la equidad y la justicia. Se sostiene que los jueces no solo aplican normas, ya que el resultado de su interpretación debe estar en consonancia con principios morales. De este modo, se defiende que «una concepción del derecho que niegue la separación absoluta entre el derecho y la moral, y que no acuda a principios de justicia material preestablecidos (como hacía el viejo iusnaturalismo) es una doctrina peligrosa» (Dworkin, 1989, p. 8). Siguiendo este planteamiento, si la IA se limita a aplicar reglas sin ponderar principios jurídicos, se podría suponer que, en las decisiones automatizadas, se crean decisiones formalmente correctas, pero que pueden ser injustas en su impacto real.

Frente a este horizonte, la respuesta compulsa a estar conscientes de que no se trata de una posibilidad lejana ni de un argumento propio de la ciencia ficción, pues hay que tener presente el valor que asume la justicia en la interpretación del Derecho. Esta transformación viene reconocida como

la integración de la tecnología digital en las distintas áreas del sistema judicial, con la finalidad de mejorar la calidad y eficacia en la resolución de conflictos, modificando tanto la organización de la justicia como la forma en que se relaciona con el ciudadano, y gestionando los riesgos generados sobre

los derechos de las personas y las garantías procesales. (Delgado, 2021, p. 28)

La implementación de la IA en la administración de justicia no debería quedar reducida a una visión simplista, constreñida en la automatización de datos y la eficiencia administrativa, sin tener en consideración el alcance que tiene y sus implicaciones en materia de DF. No se puede olvidar la razón de ser de aquella; como agente auxiliar, no posee capacidad de adaptación ni autonomía en la toma de decisiones (Castillejo, 2022, p. 64).

El rol del juez en la resolución de conflictos trasciende la mera aplicación mecánica de normas a hechos, ya que su labor implica interpretación, ponderación de principios y valoración de circunstancias específicas. Cuanto menos, resulta preocupante que, en la función jurisdiccional, el juicio humano quede relegado por decisiones automatizadas, basadas en el resultado de datos y patrones estadísticos antes que en el imperio de la ley. De ser así, cabría preguntarse cómo garantizar los DF, proteger la independencia judicial y preservar el principio de justicia. Por el momento, la respuesta es una incógnita. Por supuesto, se quiere que el debate se centre, fundamentalmente, en la implementación de unos términos que tengan como objetivo declarado el respeto de los DF y el valor que irradia la dignidad humana.

De este modo, se desea evidenciar que un sistema automatizado y basado en algoritmos carece de la capacidad para justificar una decisión en términos jurídicos y, sobre todo, en términos humanos. La robotización judicial ya es una realidad que no se puede soslayar. La posibilidad de que los algoritmos y sistemas computacionales asuman funciones que, en principio, el Derecho adjudicaba a los jueces humanos cobra, cada día, más fuerza. Esta situación se ha revertido en la aparición de jueces robots que asumen decisiones automatizadas. Por ello, resulta necesario «trabajar con la *deep learning*, integrando la base de datos con redes neuronales para entrenar a los robots jueces en el dictado de sentencias similares a las de los jueces profesionales» (Barona, 2024, p. 172).

Si se atiende al principio tradicional del Estado, como garante de los derechos en la administración de justicia, el principio de independencia cobra especial relevancia, pues, de no ser considerado, el sentido y la esencia del Derecho quedarían difuminados, ello por no hablar de la independencia como principio rector de aquella fun-

ción; si la autonomía del juez frente al sometimiento de la ley no se materializa, la decisión judicial queda vacía de legalidad. De sustituir al juez por la IA, la función jurisdiccional perdería su capacidad de tutela efectiva y se convertiría en un mero trámite, condicionado por intereses ajenos a la justicia misma (Barranco, 2005, p. 122). Esta idea obliga a tener en cuenta, de inicio, que la norma y su aplicación concreta en cada caso no es un aspecto secundario; es, en cambio, el núcleo que permite que el Derecho siga siendo un mecanismo de garantía y no un instrumento de arbitrariedad o discrecionalidad, pues la ley y la interpretación de esta ofrecen seguridad jurídica al justiciable.

Se hace forzoso, entonces, poner énfasis en la centralidad del Derecho como mecanismo de garantía. Aparte de la seguridad jurídica que se brinda al justiciable, también, cuenta cómo se concibe la justicia y su efectividad por parte del constructo social. Por lo que

el derecho es un complejo de motivos fácticos de determinación de la acción humana real que está vinculado a los demás órdenes sociales, agrupados por él en la costumbre (tradicción) y la convención (costumbre en sentido estricto). Es derecho vivo, en contraste con el derecho jurídicamente válido. (Monereo Pérez y Monereo Atienza, 2016, p. 344)

Frente a la idea de que puede optimizarse la decisión automatizada o adoptada por un sistema de IA, no se ha de pensar que el Derecho asume una proyección viva en la medida en que es observado, aceptado y reconocido por el constructo social y queda constreñido, exclusivamente, a la validez formal. En este punto, la cuestión quedaría resumida a la percepción sobre la legitimidad, ante la falta de autonomía del juez en la interpretación y aplicación del Derecho, por lo que, de ser así, podría verse afectada aquella y, por consiguiente, la probabilidad de que la sociedad acate voluntariamente las normas. Atienza —citado por García (2016)— reconoce que

hay discrecionalidad porque en Derecho no se puede decidir sin valorar, sin establecer valorativamente preferencias y sin optar entre alternativas abiertas; por ejemplo, a la hora de elegir una de entre las varias interpretaciones que de la norma son razonablemente posibles. (p. 49)

Por tanto, el Derecho no puede reducirse a un sistema determinista, en el que una norma lleve, de forma automática, a una única solución. Todo lo contrario, la discrecionalidad no se identifica, de ningún modo, con la arbitrariedad. Simplemente, responde a una exigencia valorativa entre distintas alternativas, todas ellas válidas para el Derecho. Piénsese, si no, en cómo una norma puede asumir diferentes interpretaciones en función de la órbita jurídica. La función del juez no será otra que la de aplicar una interpretación ajustada a su criterio y los valores que orientan el Derecho en una delimitación concreta.

Desde una visión constructivista, el Derecho se configura, más que como un compendio de normas, cual resultado de la interpretación y aplicación de estas por parte de los operadores jurídicos. Haciendo una extensión a lo que se aborda, se podría reconocer que el poder judicial estaría compuesto como una especie de *mente colectiva* derivada de las decisiones adoptadas por los diferentes jueces y tribunales que, con independencia de asumir decisiones de manera independiente, quedan sometidos, inexorablemente, al imperio de la ley. Podrá constatarse que

el conjunto de los integrantes del poder judicial, pueden actuar de manera para-consistente, dando soluciones diferentes para los mismos tipos de casos. Si el juez puede identificar distintos tipos de enunciados y aplicar la regla de derrotabilidad, es porque previamente se presuponen esquemas cognitivos provenientes de la teoría general del Derecho. (Martínez, 2012, p. 837)

Lógicamente, la toma de decisiones no sigue un camino único e invariable; puede adaptarse a distintos momentos y escenarios sin que, con ello, pierda la coherencia. Aplicado al Derecho, este razonamiento permite comprobar que la interpretación jurídica no ha de ser concebida como una actividad mecánica, sino como un proceso en el que se ponderan principios, se analizan las diferentes circunstancias específicas y se valora el impacto de cada decisión en la realidad social.

El juez, cual operador jurídico, en el ejercicio de sus competencias, debe resolver las diferentes controversias que se le plantean. Su obligación es interpretar y significar el hecho, con base en una norma. Para ello, deberá reconstruir la carga de la prueba y hacer una valoración racional basada en un principio de justicia.

Sobre esta premisa, si se entiende que la justicia tiene como *prima facie* la equidad y la proporcionalidad, ella se aleja, desde luego, de los sistemas algorítmicos y tecnológicos que, de entrada, pueden presentar matices discriminatorios que impiden una sentencia justa. Cuando se deposita la confianza en que una persona sea juzgada por algoritmos opacos, se desvirtúa el sentido del Derecho, bajo la justificación de la eficiencia que proporciona la IA. Si se sustituye al juez humano por el juez robot, la carga de la prueba, la testifical, los informes periciales, entre otros ejemplos, carecerían de cualquier sentido. Por ello, no se puede pensar que la actuación del juez es mecánica, como, tampoco, que su valoración se agota con la carga de la prueba de datos objetivos. La labor interpretativa del juez lleva aparejados otros aspectos de análisis, en los que la argumentación, la coherencia y la ponderación se configuran como la brújula que guía la actuación judicial. Aquí se torna más relevante la pregunta sobre la posibilidad de implementar la IA como un sistema que permita ejercer la labor judicial con las mismas garantías. Hay elementos que son difícilmente detectables por un algoritmo, como pueden ser el lenguaje corporal, la credibilidad de los testigos, la coherencia del relato de estos. Lo importante en un juicio es, verdaderamente, la valoración de las circunstancias humanas que rodean al caso concreto, pero, de admitirse la práctica de la prueba automatizada, ¿cómo se podría determinar la veracidad del testimonio? o, si existiera una duda razonable, ¿cómo se medirían los diferentes elementos en términos algorítmicos? El intento de dar una respuesta conduce a la misma conclusión: la justicia en este escenario no puede reducirse a una sentencia automatizada cuyo razonamiento sea inaccesible.

Precisamente por ello, se entiende que la decisión judicial dependerá de la valoración (que pasa por el razonamiento humano) sobre una serie de factores que influyen en la decisión y que solo un humano es capaz de interpretar y ponderar. Es más, la garantía reside en la exigencia de motivar el fallo, fundamentar el criterio y abrir la posibilidad de recurrir, en caso de que la decisión pueda ser apelada ante un tribunal superior.

Pese a todo, hay países que ya han implementado el juez robot, como China o Estonia. Es relevante que este último, desde el año 2000, cuenta con un sistema de justicia computacional. Sin embargo, fue en 2019 cuando China implementó lo que se conoce como los *Tribunales de*

internet. Un dato significativo es que ninguna de estas dos naciones ha remplazado al juez humano, pues, en ambos casos, el robot solamente se emplea en cuestiones en materia civil y las decisiones son ratificadas por el humano.

III. EL MITO DE LA NEUTRALIDAD: CUANDO EL SESGO SE ESCONDE TRAS LA OBJETIVIDAD

Hasta el momento, se han destacado los riesgos éticos que implica la implementación de la IA en la toma de decisiones judiciales frente a la racionalidad humana. En este punto, es necesario ahondar sobre los sesgos que llevan implícitos la automatización en la IA y su proyección en los derechos. En las decisiones judiciales entran en juego los DF del justiciable. Por esta razón, habrá que analizar si ellas no se ven afectadas por algoritmos discriminatorios por razón de raza, género, cultura identitaria o cualquier otro rasgo que se identifique como un factor discriminatorio. En no pocas ocasiones, se ha demostrado que la implementación de la IA en la administración de justicia, bajo la apariencia de imparcialidad, ha logrado aumentar la discriminación y perpetuarla. De hecho, cuando se ha optado por la implementación de la IA, como herramienta para evaluar los riesgos en la ejecución de las penas o la libertad condicional, los sistemas algorítmicos se presentan como soluciones, aparentemente, objetivas y eficientes. Sin embargo, cuando se señala que la IA está sesgada, rara vez se explicita qué se entiende por *sesgo* ni por qué su presencia es problemática. La ausencia de una visión crítica no hace más que invisibilizar la reproducción de desigualdades estructurales en el ámbito judicial y su consolidación.

La realidad que no se quiere ver es que la neutralidad no existe en la IA. Un algoritmo está diseñado para cumplir una función. Eso significa que se nutre de datos previos en el ámbito judicial; por lo que no va a operar en vacío. El problema se presenta cuando el algoritmo aprende y asimila datos que cuentan con patrones basados en prácticas discriminatorias. La respuesta lógica será más rigurosa en aquellos colectivos que cuentan con un patrón sesgado y discriminatorio. De esta manera, la aparente neutralidad tecnológica que se defiende con la IA quedaría sin valor, ya que, realmente, se estaría perpetuando la desigualdad (Cantero, 2024, p. 56).

Confiar en la IA para llevar a cabo esta función podría crear una falsa apariencia de independencia e imparcialidad, sin tener en cuenta que se podría estar reproduciendo una discriminación. La posición de algunos autores no se ha hecho esperar; se afirma que,

aunque se podría programar a [sic] estos sistemas para evitar sesgos basados en raza, género u otros factores discriminatorios, sigue existiendo el riesgo de que las decisiones sean influenciadas por sesgos inherentes en los datos de entrenamiento. Los algoritmos de inteligencia artificial aprenden a partir de datos históricos, y si estos datos reflejan sesgos sistemáticos, los jueces robóticos podrían perpetuar esas desigualdades. (Farfán, 2023, p. 178).

A modo de ejemplo, reflexiónese sobre lo sucedido en Inglaterra y Gales. Un error de cálculo en el formulario oficial utilizado en casos de divorcio provocó una incorrecta determinación de la pensión alimentaria en 3600 casos, a lo largo de 19 meses. El verdadero problema no radica en el fallo del sistema; cabría preguntarse por qué ni el Ministerio de Justicia ni las personas que completaron los formularios detectaron el error durante tanto tiempo (Contini, 2019, s.p.). La IA no deja de ser un sistema que se nutre del almacenamiento de datos previos. Por lo que, si cuenta con una orientación predeterminada, con base en una apariencia de neutralidad, su respuesta no siempre será imparcial y objetiva. Cotino (2019, p. 36) defiende la necesidad de una IA confiable y ética, y pone de manifiesto que la automatización no es un proceso neutral y la ausencia de un diseño responsable no haría más que reproducir y amplificar los prejuicios preexistentes.

El problema de fondo no reside en si la implementación de la IA sirve como una herramienta de auxilio en el ámbito judicial, sea justificable y beneficiosa o no, o en enrocarse sobre el modelo tradicional, sin tener en cuenta los emergentes. De lo que, verdaderamente, se trata es de que se deje al margen la capacidad crítica que tiene el juez, sin obviar que el Derecho no se presenta como un sistema de certezas absolutas, sino que se entiende más bien un entorno argumentativo, en el que la interpretación jurídica se convierte en la piedra angular de cualquier sistema de justicia. Se quiere decir con ello que no se rechazan los beneficios que genera la IA para la administración de justicia, pero, de ningún modo, se puede suplir el juicio humano sin comprometer los principios que irradia el valor de la justicia, como pueden ser la tutela,

la presunción de inocencia o la equidad. Se defiende que la implementación de la IA no ha de conceder a las máquinas el poder de decisión sobre cuestiones que afectan, directamente, los derechos y las libertades de las personas. La regulación de la IA, en sí misma, conduce a exigir, de igual manera, un marco que defina quiénes son los responsables que intervienen en el diseño y el desarrollo, puesto que, «sin poder identificar con claridad quién es responsable por los daños provocados por la actuación del juez-robot, el perjudicado quedará desamparado» (Pinerós, 2024, p. 71).

Algunos autores ponen de manifiesto que, a la hora de adoptar una decisión, el funcionamiento de la IA es similar al de los humanos, en cuanto a los sesgos y prejuicios. Es decir, cuando se identifica un sesgo algorítmico, lo que se presencia no es más que una falta de neutralidad en la respuesta, con base en los datos con los que la máquina ha sido «entrenada».

Cuando se imputa a los algoritmos de IA un sesgo, lo que se quiere decir es que los datos que arroja no son neutrales. En igualdad de circunstancias, el pertenecer a determinado grupo social, género, religión, etc., hace que la solución sea distinta. No garantizan un trato igual entre los diferentes grupos (por ejemplo, si eres de raza negra vas a prisión provisional, si eres de raza blanca a libertad provisional). (Sanchís, 2023, p. 74)

Desde luego, el debate actual se ha centrado en la implementación de la IA, con relación a los posibles sesgos discriminatorios referidos a la desigualdad de género o racial, especialmente, en escenarios en los que esta problemática está generando una vulneración de DF. De hecho, diferentes sectores aluden a la posibilidad del sesgo que hay detrás de los conocidos *algoritmos opacos*. Estos tienen un objetivo claro en su estructura de diseño predictivo. Es decir, se encargan de evaluar los niveles de riesgo y predecirlos. En este sentido se pronuncia Páez (s.f.), al advertir que,

desde el punto de vista ético, plantea muchas de las mismas preguntas que la opacidad jurídica con respecto a la discriminación oculta y la violación de los derechos humanos. Desde el punto técnico, es un obstáculo para los desarrolladores que quieran mejorar el desempeño del modelo, y detectar y resolver sesgos y otros riesgos semejantes. (s.p.)

Se concuerda en la doble dimensión en la que se proyecta el problema de los algoritmos opacos, puesto que, por un lado, se dimensionan por tratarse de una clara vulneración de los DF y, por la otra, es notable la falta de transparencia técnica.

La lógica de funcionamiento, verdaderamente, despierta las alarmas, sobre todo, porque la estructura y metodología del diseño de los sistemas automatizados quedan amparadas por el Derecho de propiedad intelectual de los creadores; por lo que se evidencia una falta de transparencia que impide que las personas afectadas conozcan los criterios exactos que determinan las decisiones que les conciernen (Roa, 2022, p. 278).

Los programas de IA son creados y desarrollados, en la mayoría de los casos, por entidades privadas y se amparan en el Derecho sobre la propiedad intelectual que garantiza la inversión tecnológica. Ello lleva implícita la imposibilidad de conocer la estructura del sistema y los datos que se han incluido para su desarrollo.

Si en la fase procesal, al justiciable le resulta incomprensible el fallo, el derecho a la tutela judicial efectiva o a un juicio justo, como lógica respuesta, tiene un sistema de justicia carente de legitimidad. En atención a lo anterior, se defiende que la opacidad y la falta de transparencia atentan, directamente, contra el derecho a la tutela judicial. ¿Cómo puede un ciudadano ejercer su derecho a la defensa, si no sabe qué criterios se han utilizado para valorar su caso? ¿Cómo puede un juez verificar la equidad de una decisión, si el sistema que la ha generado no es auditable ni cuestionable?

Hasta el momento, no hay respuesta. Bajo este planteamiento, habrá que pensar en la reformulación de un derecho que garantice la tutela efectiva y el derecho del justiciable a estar informado sobre cualquier decisión judicial que lo afecte. En principio, esto podría parecer algo descabellado; sin embargo, es una cuestión de primer orden en la pretensión de implementar la automatización en las decisiones judiciales. La automatización de estas no se puede alcanzar a cualquier costo, sobre todo cuando entra en juego la vulneración de los DF. En el prisma del sentido de la justicia, no cabe la posibilidad de que la decisión sea incomprensible por el justiciable. Si se reconoce al «lenguaje judicial como [aquel] emplead[o] por los diferentes entes encargados de administrar justicia en sus relaciones

con los ciudadanos» (Apa, 2021, p. 151), se hace necesario que él sea claro e inteligible para el justiciable. De lo contrario, el derecho a la defensa quedaría como un muro infranqueable en el acceso real a los derechos, y minaría la legitimidad y eficacia del sentido del Derecho, por lo cual se evidencia un significativo riesgo del derecho a la defensa, la tutela judicial y el derecho a una sentencia justa y sin dilaciones indebidas.

Es preocupante verificar, ciertamente, hasta qué punto la implementación de la IA en la administración de justicia tiene capacidad para preservar los valores que fundamentan las funciones del Derecho, como la equidad, la proporcionalidad y el respeto a los DF. Desde luego, esto resulta aberrante y un despropósito en clave de garantía procesal; desde una visión de justicia, los jueces y tribunales deben ser imparciales; su actuación, también, ha de ser verificable y comprensible para quienes se ven afectados por sus decisiones. La opacidad no puede tratarse como un simple defecto técnico; más bien debe asumirse como lo que, realmente, significa: una vulneración del derecho a la defensa en materia de DF, pues, si el justiciable no puede cuestionar la base de su propia condena, ¿cómo puede hablarse de un juicio justo?

Un ejemplo palmario se encuentra en COMPAS que, traducido del inglés, significa «Administración de perfiles de criminales para sanciones alternativas del sistema de prisiones de los EE.UU.» y ha sido implementado por varios Estados norteamericanos, con el objetivo de evaluar el riesgo de reincidencia de los penados (Miró, 2018, p. 108). Sobre la justificación de ofrecer rigor y previsibilidad en la toma de decisiones judiciales, numerosos estudios muestran cómo la implementación de ese sistema ha generado un mayor impacto en la raza afroamericana, en comparación con las personas blancas, a pesar de que contaban con los mismos perfiles delictivos de riesgo. Es reseñable, en tal sentido, el estudio realizado por ProPublica (Larson, Mattu y Angwin, 2016, s.p.) en el que identificaron que COMPAS situaba el factor del colectivo afroamericano con un mayor porcentaje de reincidencia (44,9%), frente a los acusados de raza blanca (23,5%). De hecho, a la inversa, en el estudio sobre la baja reincidencia, las personas blancas contaban con una tasa del 47,7%, frente a los afroamericanos que disminuía en un 28%.

Esto, desde luego, evidencia un modelo de justicia que, lejos de corregir las desigualdades existentes en el sistema penal, parece perpetuarlas, bajo la apariencia de una supuesta objetividad matemática y, a la vez, que se reproducen desigualdades estructurales y se eternizan, como si se tratara de una legitimidad renovada. Ello conduce a reflexionar sobre el principio de responsabilidad individual, conforme con el cual las personas deben ser juzgadas por sus actos, mas no por la probabilidad estadística de que reincidan en el futuro.

Al igual que en el caso de COMPAS, existen otros programas como el *ShotSpotter*, creado para la detección temprana de posibles tiroteos, mediante la alerta a la policía. De igual forma, este sistema ha sido cuestionado por sus sesgos y la falta de transparencia en las decisiones judiciales. Prueba de ello es el estudio realizado por MacArthur Justice Center, en el que se hizo un diagnóstico cuantitativo de las estaciones de policía en Chicago. Los datos revelaron que cerca del 89% de las alertas generadas por *ShotSpotter* no resultaron en la identificación de crímenes relacionados con armas de fuego y el 86% no identificaron ningún tipo de crimen (Associated Press, 2023).

Lo expuesto pone en evidencia la falacia de la neutralidad algorítmica. Desde una visión ética, pensar que los datos automatizados puedan proveer respuestas desprovistas de valoraciones ideológicas ignora el rol de quienes diseñan y desarrollan la tecnología. Se olvidan, quienes así piensan, que la realidad es que cualquier sistema de IA es el resultado de una voluntad humana y, en función de su diseño, la toma de decisión tendrá en cuenta las variables (datos históricos, precedentes judiciales, etc.). Vistas las cosas así, la implementación tecnocrática en la administración de justicia no hace más que reducir al individuo a una serie de datos cuantificables. Tratándose de DF, al menos resulta preocupante encomendar la responsabilidad a un sistema opaco y discriminatorio. Esto se podría entender, casi como una renuncia a la responsabilidad que el Derecho tiene de garantizar la dignidad y la igualdad de todas las personas ante la ley, sin prejuicio de raza, género o cualquier otra diferenciación.

A mayor abundamiento, si se acepta y normaliza la opacidad de esos sistemas, sin cuestionar su legalidad, no se hace más que renunciar a los principios que inspiran el Estado de Derecho. En este sentido, se pronuncia Monasterio (2017), al reconocer que «garantizar todos los derechos de las personas y en particular aquellos que se pueden

conculcar a partir de los usos innovadores de los datos personales no es algo abstracto y difícil de conseguir» (p. 214).

En la misma línea, otra cuestión que está presente en el debate sobre la implementación de la IA en el sistema de justicia es la que afecta la responsabilidad. Moralmente, las decisiones automatizadas carecen de responsabilidad, ya que no cuentan con conciencia ni ética. Es alarmante cómo se desdibuja la responsabilidad del operador jurídico, bajo el manto de la supuesta objetividad que ofrece la IA para eludir sus obligaciones éticas y jurídicas en el ejercicio de la función social que representa.

Ya se ha advertido que, entre los riesgos más temibles, están, precisamente, los algoritmos que cuentan, por un lado, con la posibilidad de tener un sesgo discriminatorio y, del otro, la opacidad y falta de transparencia que se enarbola a partir del derecho de propiedad intelectual de sus creadores. Son varios los interrogantes que no cuentan con una fácil respuesta o, por lo menos, hasta el momento, esta no resulta del todo significativa, al no quedar claro sobre quién recae la responsabilidad, cuando el algoritmo cuenta con un sesgo discriminatorio: si sobre el programador que diseñó el modelo, quienes diseñaron el programa de *software* o el juez que adoptó una decisión generada por la IA. Sobre este parecer, algunos autores reconocen que

los humanos serían los «autores intelectuales» de las acciones de una IA. Por otro lado, puede atribuirse responsabilidad a la persona que dirige o es propietaria de un sistema de IA, de la misma manera que el propietario de una empresa es responsable por las acciones de esta, aunque no las ordene directamente. (Espínosa, 2023, p. 6)

IV. CONCLUSIONES

La implementación de la IA y los sistemas automatizados no pueden asumirse con un entusiasmo ciego, como, tampoco, con la ingenua creencia de que, por sí sola, esta puede alcanzar la eficiencia de la administración de justicia. Ante la cuestión, sumamente importante, de si la IA puede juzgar, o no, se defiende que la justicia no es un mero ejercicio de predictibilidad estadística ni de procesamiento de datos, sino un acto de ponderación de valores, de interpretación de principios

y de reconocimiento que solo el humano puede realizar. La justicia no se reduce a una neutralidad matemática, ya que se fundamenta en la legitimidad democrática.

El riesgo de trasladar estas decisiones a un sistema algorítmico expone a que las decisiones provengan de sistemas opacos. El peligro latente recae sobre quién controla la IA en la justicia. Los algoritmos son diseñados, programados y gestionados por corporaciones privadas o estatales, y no siempre el objetivo es la justicia; se busca, antes, la eficiencia, el control e, incluso, la rentabilidad económica. De igual manera, habría que detenerse en los sesgos que han demostrado las numerosas investigaciones realizadas hasta el momento.

La IA puede ser una herramienta de apoyo en la administración de justicia, y debe serlo, pero nunca un sustituto del juez humano. Se trata de aceptar la tecnología y su utilidad, pero con plena comprensión de sus límites. Si se confía, ciegamente, en la IA para impartir justicia, se asistiría a un mundo en el que las decisiones judiciales no serían actos de justicia, sino simples cálculos impávidos y opacos; por no decir que la justicia, sin humanidad, no es justicia; es solo un espejismo de equidad, una simulación de legalidad, una sombra de lo que debería ser. Es aquí donde reside la mayor responsabilidad de los juristas y teóricos del Derecho: en garantizar que el progreso tecnológico no haga olvidar lo esencialmente humano de la justicia. Al final, la pregunta no es si la IA puede juzgar; primero, hay que inquirir qué tipo de justicia se desea para la sociedad.

V. REFERENCIAS

- Alastruey Merino, C. F. (2021). Estado de la cuestión de la inteligencia artificial y los sistemas de aprendizaje autónomo. *Sociología y Tecnociencia*, 11(2), 182-195. https://doi.org/10.24197/st.Extra_2.2021-182-195
- Apa, M. J. (2021). El lenguaje judicial y el derecho a comprender. *Pensar en Derecho*, 18, 149-177. <http://www.derecho.uba.ar/publicaciones/pensar-en-derecho/revistas/18/el-lenguaje-judicial-y-el-derecho-a-comprender.pdf>
- Associated Press. (Enero 20, 2023). *Confidential document reveals key human role in gunshot tech*. <https://apnews.com/article/>

[shotspotter-artificial-intelligence-investigation-9cb47bbfb565dc3ef110f92ac7f83862](https://doi.org/10.2436-rdcp.i58.2019.3303)

- Barona Vilar, S. (2024). El dilema del razonamiento, motivación y decisión judicial algorítmica (IA): ¿La posibilidad de lo imposible? *Teoría & Derecho*, 37, 150-175. <https://doi.org/10.36151/TD.2024.109>
- Barranco Avilés, M. C. (2005). El papel del juez en el Estado constitucional. *Anuario de Filosofía del Derecho*, 22, 113-130. <https://revistas.mjusticia.gob.es/index.php/AFD/article/view/2096>
- Cabrera Fernández, M. (2024). Aplicación de la inteligencia artificial a la toma de decisiones judiciales. *Eunomía*, 27, 183-200. <https://doi.org/10.20318/eunomia.2024.9006>
- Cantero Gamito, M. (2024). Acceso a la justicia en tiempos de IA: ¿hacia una justicia low-cost? *CIDOB*, 138, 51-71. <https://www.raco.cat/index.php/RevistaCIDOB/article/view/433783>
- Castillejo Manzanares, R. (2022). Digitalización y/o inteligencia artificial. En Calaza López, S. y Sánchez Arjona, M. (Dtores.). *Inteligencia artificial legal y administración de justicia*. Thomson-Reuters-Aranzadi.
- Contini, F. (s.f.). La inteligencia artificial: ¿un nuevo caballo de Troya que influye indebidamente en el poder judicial? *UNODC*. <https://www.unodc.org/dohadeclaration/es/news/2019/06/artificial-intelligence-a-new-trojan-horse-for-undue-influence-on-judiciaries.html>
- Cotino Hueso, L. (2017). *Big data* e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales. *Dilemata*, 24, 131-150. <https://www.dilemata.net/revista/index.php/dilemata/article/view/412000104>
- Cotino Hueso, L. (2019). Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y *big data* confiables y su utilidad desde el Derecho. *Revista Catalana de Dret Públic*, 58, 29-48. <https://revistes.eapc.gencat.cat/index.php/rcdp/article/view/10.2436-rdcp.i58.2019.3303>

- Delgado Martín, J. (2021). Reflexiones sobre el estado actual de la transformación digital de la justicia. *Acta Judicial*, 8, 27-43. <https://revistaactajudicial.letradosdejusticia.es/index.php/rajj/article/view/58>
- Dworkin, R. (1989). *Los derechos en serio* (2.^a ed.). Ariel.
- Espinosa, P. y Clemente, M. (2023). La percepción de la toma de decisiones a través de inteligencia artificial cuando se produce daño a las personas. *Estudios Penales y Criminológicos*, 44, 1-13. <https://revistas.usc.gal/index.php/epc/article/view/8917>
- Farfán Intriago, J. L., Farfán Largacha, J. A., Farfán Largacha, B. y Núñez Vera, J. P. (2023). Inteligencia artificial y Derecho: ¿La justicia en manos de la IA? *Frónesis*, 30(2), 173-197. <https://produccioncientificaluz.org/index.php/fronesis/article/view/40853>
- Ferrajoli, L. (2006). *Sobre los derechos fundamentales y sus garantías* [Carbonel, M., Cabo, A. y Pisarello, G. (Trad.)]. Comisión Nacional de Derechos Humanos de México.
- García Amado, J. A. (2016). ¿Para qué sirve la teoría de la argumentación jurídica? *Teoría & Derecho*, 20, 22-49. <https://ojs.tirant.com/index.php/teoria-y-derecho/article/view/455>
- García-Marzá, D. (2024). Crítica de la razón algorítmica: contra la neutralidad como ideología. En González-Esteban, E. y Siurana Aparisi, J. (Eds.). *Inteligencia artificial: concepto, alcance, retos*, 106. Tirant Humanidades.
- Habermas, J. (1984). *Ciencia y técnica como ideología* [Jiménez Redondo, M. (Trad.)]. Tecnos.
- Herbert Hart, L. A. (1961). *El concepto de derecho* [Carrió, G. R. (trad.)]. Abeledo-Perrot.
- Larson, J., Mattu, S. y Angwin, J. (s.f.). Cómo analizamos el algoritmo de reincidencia COMPAS. *ProPública*. <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>
- Martínez Bahena, G. C. (Septiembre-diciembre, 2012). La inteligencia artificial y su aplicación al campo del Derecho.

- Alegatos*, 82, 827-846. <https://alegatos.azc.uam.mx/index.php/ra/article/view/205>
- Miró Llinares, F. (2018). Inteligencia artificial y justicia penal: más allá de los resultados lesivos causados por robots. *Revista de Derecho Penal y Criminología*, 20, 107-130. <https://doi.org/10.5944/rdpc.20.2018.26446>
- Monasterio Astobiza, A. (2017). Ética algorítmica: Implicaciones éticas de una sociedad cada vez más gobernada por algoritmos. *Dilemata*, 24, 185-217. <https://dilemata.net/revista/index.php/dilemata/article/view/412000107>
- Monereo Pérez, J. L. y Monereo Atienza, C. (2016). *Sociología jurídica en Max Weber: economía, sociedad y derecho*, 344. Centro de Investigación y Docencia Económicas (CIDE). <https://www.econbiz.de/Record/sociolog%C3%ADa-jur%C3%ADdica-en-max-weber-econom%C3%ADa-sociedad-y-derecho-monereo-p%C3%A9rez-jos%C3%A9-luis/10011588392>
- Otero Parga, M. (2023). ¿Puede la inteligencia artificial sustituir a la mente humana? Implicaciones de la IA en los derechos fundamentales y en la ética. *Anales de la Cátedra Francisco Suárez*, 57, 39-61. <https://doi.org/10.30827/acfs.v57i.24710>
- Pazos Sierra, J. y García Codina, C. (2024). La inteligencia artificial general y su llegada. *I+S*, 158, 7-13. <https://seis.es/wp-content/uploads/2024/03/IS-158.pdf#page=7>
- Pineros Polo, E. (2024). El juez-robot y su encaje en la Constitución española: La inteligencia artificial utilizada en el ámbito de la toma de decisiones por los tribunales. *Estudios de Deusto*, 72(1), 53-78. <https://revista-estudios.revistas.deusto.es/article/view/3100>
- Roa Avella, M. P., Sanabria-Moyano, J. E., Dinas-Hurtado, K. (2022). Uso del algoritmo COMPAS en el proceso penal y los riesgos a los derechos humanos. *Revista Brasileira de Direito Processual Penal*, 8(1), 275-310. <https://doi.org/10.22197/rbdpp.v8i1.615>
- Rouhiainen, L. (2018). *Inteligencia artificial*. Alienta Editorial.

- Sanchís Crespo, C. (2023). Inteligencia artificial y decisiones judiciales: crónica de una transformación anunciada. *Scire, Representación y Organización del Conocimiento*, 29(2), 65-84. <https://roderic.uv.es/items/bc1e94a5-b829-4815-a738-1d54943d3659>
- Truyol Serra, A. (1991). Bases filosóficas y metodológicas para un derecho de la sociedad de la información. En *Implicaciones socio-jurídicas de las tecnologías de la información*, 118. Centro de Informática, Telemática y Medios Afines.